JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 32, 505–511 (1970)

On Factoring Positive Operators

J. RISSANEN

IBM Research Division, San Jose, California

Submitted by L. A. Zadeh

The problem of factoring positive operators into an "outer" factor and its adjoint has been studied, and its application to the theory of non-stationary random processes is discussed. This problem is an extension of the classical one of factoring positive matrix-valued functions defined on the unit circle of the complex plane into similar factors.

1. INTRODUCTION

The problem of factoring positive matrix-valued functions defined on the unit circle of the complex plane into an "outer" factor and its adjoint has been the subject of numerous investigations, [1–8]. The importance of this problem derives from its close and natural relationship to the theory of stationary multivariate random processes. Such a process was shown to admit a one-sided moving average representation, or, equivalently, to be non-deterministic precisely when the spectral measure of its covariance function factors in the indicated manner.

We describe briefly this factorization in a slightly more abstract form [8]. Let H be a Hilbert space and $V : H \to H$ an isometric so-called shift operator such that H admits the decomposition, $H = \sum_{i\geqslant 0} V^i L$, where $L = (VH)^\perp$. An operator A is called "Toeplitz" if $V^*AV = A$, "analytic" if $VA = AV$, and "outer" or "optimal" if it is analytic and its range is dense in a subspace which reduces V. The factorization problem is to represent a positive Toeplitz operator as A^*A, where A is outer or, alternately, analytic.

Although the shift operator appears to be central in the foregoing definitions, it can be removed without any essential loss in the meaning of these notions. Our plan is to study how this is done in such a manner that a factorization of non-Toeplitz operators becomes possible. Moreover, the new factorization bears the same close relationship to—this time—nonstationary random processes as the old does in the case of stationary processes.

Specifically, we determine both necessary and sufficient conditions for positive definite symmetric operators to factor as A^*A with A being "analytic" or "outer" in a more general sense. Thereafter we show that the same conditions characterize nonstationary random processes to admit a one-sided moving average representation, or, equivalently, to be purely non-deterministic [10].

2. Analytic and Outer Factors

If $H = H_0 \supset H_1 \supset H_2 \cdots$ is a chain of subspaces of a Hilbert space H, then H decomposes as a direct sum of orthogonal subspaces in the following way (a Wold-decomposition):

$$H = \sum_{i\geqslant 0} L_i \oplus \bigcap_{i\geqslant 0} H_i, \tag{1}$$

where $L_i = H_{i+1} \cap H_i^{\perp}$. To see this, observe that $H = H_0 = \sum_{i=0}^{n-1} L_i \oplus H_n$, and that an element x belongs to the orthogonal complement of $\sum_{i\geqslant 0} L_i$ iff it belongs to each of the spaces H_n.

When the subspaces H_n are obtained by an isometric operator V as $H_n = V^n H$, then the spaces L_i become "wandering" [9], i.e., $L_i = V^i L_0$, and V on these acts like a shift. Moreover, $\bigcap_{i\geqslant 0} H_i$ reduces V.

In general, no mapping from a space L_i onto its successor L_{i+1} need exist, but that causes no particular difficulty in the present study.

From now on, our basic Hilbert space in which the various operators act is separable and admits a Wold-decomposition in which

$$\bigcap_{i=0}^{\infty} H_i = \{0\}; \qquad \text{i.e.,}$$

$$H = \sum_{i=0}^{\infty} L_i \tag{2}$$

Let (x, y) denote the inner product and $\| x \|$ the norm; also write

$$H_n = \sum_{i=n}^{\infty} L_i .$$

Let $Q : H \to H$ be a positive definite symmetric operator with domain D_Q consisting of all the elements of the form $x = x_0 + \cdots + x_n$, $x_i \in L_i$, some n. In other words, $(x, Qy) = (Qx, y)$ for all $x, y \in D_Q$, and $(x, Qx) = 0$ only if $x = 0$. The last condition for positive definiteness is made for convenience; it is easy to relax.

We call an operator $A : H \to H$ *analytic* if

(i) Domain D_A contains the subspaces L_i

(ii) $A : H_n \to H_n$ for all n.

Let K_n denote the completion of the range of an analytic operator A when restricted to H_n. By (ii), $K_n \subseteq H_n$. We call an analytic operator A *outer* if:

(iii) $N_n \triangleq K_n \cap K_{n+1}^{\perp} \subseteq L_n$ for all n.

The problem we shall be concerned with is to find out when Q factors as

$$Q = A^*A,\ D_A = D_Q,$$

with A analytic or, alternately, outer.

3. Factorization Theorem

The quadratic form (x, Qy) induces a second inner product, denoted $\langle x, y\rangle$, in the linear manifold spanned by the L_i's. Let $\hat{H}$ denote the completion of this manifold with respect to the metric induced by the inner product $\langle x, y\rangle$. Similarly, let $\hat{H}_n$ denote the subspace spanned by the spaces L_i for $i \geqslant n$. With $\hat{L}_i = \hat{H}_i \cap \hat{H}_{i+1}^{\perp Q}$, where the letter Q emphasizes the orthogonality in $\hat{H}$, the decomposition (1) gives

$$\hat{H} = \sum_{i=0}^{\infty} \hat{L}_i \oplus_Q \bigcap_{i=0}^{\infty} \hat{H}_i. \tag{3}$$

As opposed to (2), the subspace $\bigcap_{i=0}^{\infty} \hat{H}_i$ may no longer be zero; its dimensionality depends clearly on the positive operator Q.

(4) Lemma. $\dim \hat{L}_n \leqslant \dim L_n$

Proof. Let P_n denote the orthogonal projection of $\hat{H}$ onto $\hat{L}_n$. Then, P_n maps $L_n(\subset \hat{H})$ into a dense set in $\hat{L}_n$. To see this, let $x \in \hat{L}_n$ and x be Q-orthogonal (i.e., orthogonal in the inner product of $\hat{H}$) to $P_n(L_n)$. Then, for any $y \in L_n$, $\langle x, y\rangle = \langle P_n x, y\rangle = \langle x, P_n y\rangle = 0$; i.e., x is Q-orthogonal to L_n. Further, since x being in $\hat{L}_n$ is Q-orthogonal to $\hat{H}_{n+1}$, x is Q-orthogonal to the dense set $L_n + \hat{H}_{n+1}$ in $\hat{H}_n$ and, hence, $x = 0$.

This proves the lemma if $\dim L_n$ is finite; the countable case follows from the separability of the spaces involved.

Our main theorem is the following, which extends a similar result by Lowdenslager [1].

(5) THEOREM. *A factorization $Q = A^*A$ with $D_A = D_Q$ and A analytic or outer exists iff $\bigcap_{i\geqslant 0} \hat{H}_i = \{0\}$.*

Proof. Let $\bigcap_{i=0}^{\infty} \hat{H}_i = \{0\}$. By Lemma (4) we may select a subspace $N_n \subseteq L_n$ with $\dim N_n = \dim \hat{L}_n$ and an isometric mapping $G_n : \hat{L}_n \to N_n$ preserving the inner products, i.e., $(G_n x, G_n y) = \langle x, y\rangle$ for all $x, y \in \hat{L}_n$. Define $G : \hat{H} \to H$ by: $Gx = \sum G_i x_i$, where $x = \sum x_i$, $x_i \in \hat{L}_i$. Clearly, $(Gx, Gy) = \langle x, y \rangle$ for all $x, y \in \hat{H}$.

Define $A : H \to H$ as the restriction of G on the set D_Q. Then for all $x, y \in D_A$, $\langle y, x\rangle = (y, Qx) = (Ay, Ax)$, which means that $Qx = A^*Ax$, i.e., $Q = A^*A$. We show next that $A : H_n \to H_n$ and, hence, that A is analytic. This is true since G, mapping $\hat{H}_n$ into H_n, maps $L_n \subset \hat{H}_n$ into H_n, and so does A. Condition (iii) holds by the construction which shows that A is outer.

Conversely, let $Q = A^*A$, with $D_A = D_Q$ and A analytic. If $x \in D_A$ then $Ax \in D_{A^*}$ by the assumed factorization, and $(Ax, Ax) = (x, A^*Ax) = \langle x, x\rangle$. Hence A maps isometrically a dense submanifold of $\hat{H}$ into H. Extend A to an isometric mapping G defined on all of $\hat{H}$. Clearly G maps $\hat{H}_n$ into H_n. If $x \in \bigcap_{n\geqslant 0} \hat{H}_n$ then $Gx \in \bigcap_{n\geqslant 0} H_n = \{0\}$. Since G is an isometry, $x = 0$, which proves the converse of the theorem.

The reasoning in the proof of the converse part of the preceding theorem provides a fairly simple sufficient condition for Q to factor.

(6) COROLLARY. *If $\hat{H} = \{x \in H \mid (x, Qx) < \infty\}^C \subseteq H$, where the completion is with respect to the metric induced by (x, Qy), then the positive definite symmetric operator Q factors.*

Proof. Let $\phi : \hat{H} \to H$ denote the identity map which embeds $\hat{H}$ into H. ϕ maps $\hat{H}_n$ into H_n, and hence $\bigcap_{n=0}^{\infty} \hat{H}_n$ into $\bigcap_{n=0}^{\infty} H_n = \{0\}$. Since ϕ is one-to-one, $\bigcap_{n=0}^{\infty} \hat{H}_n = \{0\}$, and Theorem (5) applies.

In particular, by Corollary (6) any operator Q which satisfies $k_1 e \leqslant Q \leqslant k_2 e$, where e is the identity, factors [11]. But also if the manifold $\{x \in H \mid \| x \|_Q < \infty\}$ is complete, we have a case where Q factors. Here, $\| x \|_Q = (x, Qx)$ if $x \in D_Q$ and $\| x \|_Q = \lim(x_n, Qx_n)$ for all $x \in H$ such that the limit exists when $x_n \to x$, $x_n \in D_Q$.

4. APPLICATION TO RANDOM PROCESSES

Consider a random process $y = \{y_n\}_{n\geqslant 0}$, where at each integer-valued time instance n (counted backwards), y_n is a family $(y_n^0, y_n^1, \ldots)$ of random variables y_n^i considered as elements in a Hilbert space. Such a process is said to have a one-sided moving average if each y_n^i can be written as

$$y_n^i = \sum_{m \geqslant n} \sum_{j \geqslant 0} C_{nm}^{ij} x_m^j, \tag{7}$$

where 1) the process $x = \{x_n\}$ with $x_n = (x_n^0, x_n^1, \ldots)$ is orthonormal, i.e., $(x_n^i, x_m^j) = 0$ if $i \neq j$ or $n \neq m$, and $\| x_n^i \| = 1$ for all n and i; and, 2) the prediction error $\| y_n^i - \hat{y}_n^i \|^2 = \sum_{j \geqslant 0} (C_{nn}^{ij})^2$, with $\hat{y}_n^i$ denoting the orthogonal projection of y_n^i onto the subspace Y_{n+1} spanned by the elements $\{ y_m^j \mid m > n, j \geqslant 0\}$.

These conditions insure that the projection admits the expansion

$$\hat{y}_n^i = \sum_{m > n} \sum_{j \geqslant 0} C_{nm}^{ij} x_m^j. \tag{8}$$

In fact, since Y_{n+1} is contained in the subspace X_{n+1} spanned by $\{x_m^j \mid m > n, j \geqslant 0\}$, the predicion error is no smaller than the given number $\sum_{j \geqslant 0} (C_{nn}^{ij})^2$, which by (7) is the squared distance of y_n^i to X_{n+1}. With equality the expansion (8) must hold since the projection is unique.

The one-sided moving average representation (7) is central in the theory of prediction. For one thing, the formula (8) provides a partial solution to the prediction problem. (The remaining problem is to express the x-variables in (8) in terms of the y-variables whenever this is possible.) But also, the existence of such a representation is tantamount to the process of being "purely nondeterministic" [10].

Our plan is to express the existence of the moving average representation in terms of the covariance function. We also show how the notion of an outer operator gets a natural interpretation in much the same way as was done in [5] and [7] in the case of stationary processes. More precisely, we show that a moving average representation exists if and only if the covariance function, interpreted as a symmetric positive operator in a Hilbert space, factors as A^*A with A outer.

We shall have to restrict the class of processes considered by requiring that the covariance matrices $Q_{nn} = \{Q_{nn}^{ij} \mid i, j \geqslant 0\}$, where $Q_{nm}^{ij} = (y_n^i, y_m^j)$, be bounded; for each set of numbers $a_0, a_1, \ldots, a_N$ such that $\sum a_i^2 = 1$, let $\sum_{i,j} a_i Q_{nn}^{ij} a_j < M$, where M might depend on n This condition is certainly satisfied in the perhaps most interesting case where each component y_n of the process has only finitely many random variables.

Suppose, first, that a moving average representation (7) for a process y exists. For each n, let L_n denote the space spanned by the orthonormal elements $x_n^0, x_n^1, \ldots$, and form $H = \sum_{i \geqslant 0} L_i$.

Let the matrix $\{C_{nm}^{ij}\}$ with $C_{nm}^{ij} = 0$ for $n > m$ define a mapping C on the set $\{x_n^i\}$ by

$$Cx_n^i = \sum_{m \geqslant n} \sum_{j \geqslant 0} C_{nm}^{ij} x_m^j.$$

Extend C by linearity to the span of the set $\{x_n^i\}$. The adjoint of C, C^*, is then given by

$$C^* x_n^i = \sum_{m=0}^{n} \sum_{j \geqslant 0} C_{mn}^{ji} x_m^j,$$

which makes the operator $Q = C^*C$ symmetric. We may identify Q with the covariance function written as the matrix $\{Q_{nm}^{ij}\}$. By the boundedness assumption of the matrices Q_{nn}, C and Q may be extended once more to include each of the subspaces L_i in their domain (condition (*i*) of analytic operators).

The mapping C is, then, analytic. The completion of the range of C, restricted to $H_n = \sum_{i \geqslant n} L_i$, is the linear subspace Y_n spanned by the vectors $\{ y_m^i \mid i \geqslant 0, m \geqslant n\}$. The subspace $N_n = Y_n \cap Y_{n+1}^{\perp}$ is spanned by the vectors $\{ y_n^i - \hat{y}_n^i \mid i \geqslant 0\}$ where $\hat{y}_n^i$ is the orthogonal projection of y_n^i onto Y_{n+1}. By (7) and (8) we conclude that N_n is spanned by the vectors $\{\sum_{j \geqslant 0} C_{nn}^{ij} x_n^j \mid i \geqslant 0\}$, which finally implies that $N_n \subseteq L_n$, and hence, that C is outer.

Conversely, we assume that there exists a space $H = \sum_{i \geqslant 0} L_i$ with an orthonormal basis $\{x_n^i \mid i \geqslant 0\}$ spanning L_n for each n, with respect to which the covariance matrix $Q = \{Q_{nm}^{ij}\}$ factors as A^*A, where $A = \{A_{nm}^{ij}\}$ is analytic. By theorem (5) we may assume that A is outer. Further, by rotating the spanning set for L_n by a unitary operator, if necessary, we may furthermore assume that some subset of these vectors spans $N_n \subseteq L_n$, where N_n is defined by condition (iii) for outer mappings.

We wish to construct a process y from certain elements of H such that $(y_n^i, y_m^j) = Q_{nm}^{ij}$, and a one-sided moving average representation for y.

Define the process y by

$$y_n^i = Ax_n^i, i \geqslant 0, n \geqslant 0 \tag{9}$$

Then, its covariance function is the matrix Q, for

$$(y_n^i, y_m^j) = (Ax_n^i, Ax_m^j) = (x_n^i, Qx_m^j) = Q_{nm}^{ij}.$$

Plainly, y_n^i lies in the space $K_n = N_n \oplus N_{n+1} \cdots$. Consequently, it admits the expansion

$$y_n^i = \sum_{m \geqslant n} \sum_{j \geqslant 0} C_{nm}^{ij} x_m^j, \tag{10}$$

where the coefficients are given by $C_{nm}^{ij} = (y_n{}^i, x_m{}^j) = (Ax_n{}^i, x_m{}^j) = A_{mn}^{ji}$. Observe that $C_{nm}^{ij} = 0$ if $m < n$ or if $x_m{}^j \perp N_m$.

Equation (10) is a one-sided moving average representation for the process y. To see this, observe first that $\| y_n{}^i - \hat{y}_n{}^i \| \geqslant \sum_{j \geqslant 0} (C_{nn}^{ij})^2$ by the same arguments as above following (8). The equality follows from the fact that the elements $y_m{}^j$, $m \geqslant n$, and $j \geqslant 0$ span the space K_n.

References

1. D. Lowdenslager, On factoring matrix valued functions, *Ann. of Math.* **78** (November, 1963), 450–454.
2. A. Devinatz, The factorization of operator valued functions, *Ann. of Math.* **73** (1961), 458–495.
3. A. Beurling, On two problems concerning linear transformations in Hilbert space, *Acta Math.* **81** (1949), 239–255.
4. N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, *Acta Math.* **98** (1957), 111–150; **99** (1958), 93–137.
5. P. Masani, Shift invariant spaces and prediction theory, *Acta Math.* **107** (1962), 275–290.
6. H. Helson, "Lectures on Invariant Subspaces," Academic Press, 1964.
7. H. Helson and D. Lowdenslager, Vector-valued processes, *Proc. Fourth Berkeley Symposium* **II** (1961), 203–212.
8. M. Rosenblum, Vectorial Toeplitz operators and the Fejer–Riesz theorem, *J. Math. Anal. Appl.* **23** (1968), 139–147.
9. P. Halmos, Shifts on Hilbert spaces, *Crelle's J.* **208** (1961), 102–112.
10. H. Cramer, On some classes of non stationary stochastic processes, *Proc. Fourth Berkeley Symposium* **II** (1961), 57–78.
11. J. Rissanen and L. Barbosa, A factorization problem and the problem of predicting non stationary vector-valued stochastic processes, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **12** (1969), 255–266.